Εργασία [Ανάλυσης Κοινωνικών Δικτύων](http://gunet2.cs.unipi.gr/courses/TMC132/)

**Ιωακείμ Ελ-Χαττάμπ-Μπριστογιάννης Π19048**

Η παρακάτω εργασία υλοποιήθηκε στο προγραμματιστικό περιβάλλον **MATLAB.**

Αρχικά , διαβάζουμε το αρχείο sx-stackoverflow.txt με την σωστή μορφή και αποθηκεύουμε τα δεδομένα στην μήτρα **data** και στην συνέχεια επιλέγουμε πόσες χρονικές στιγμές επιθυμούμε να έχει η ανάλυση καθώς και τις κορυφαίες τιμές ομοιότητας.

##### Ερώτημα 1ο

Χρησιμοποιώντας της ενσωματωμένες συναρτήσεις min και max βρίσκουμε την αρχική και τελική χρονική στιγμή του συνόλου δεδομένων μας.

##### Ερώτημα 2ο

Στην συνέχεια χρησιμοποιώντας τους τύπους:

για

Υπολογίζουμε τις χρονικές περιόδους της ανάλυσης

##### Ερώτημα 3ο

Με την βοήθεια του αντικειμένου **digraph** του **MATLAB** αναπαριστούμε τα κατευθυνόμενα υποδίκτυα των χρονικών στιγμών, κοιτάζοντας την κάθε ακμή στα δεδομένα και τοποθετώντας την στο κατάλληλο υποδίκτυο με χρήση της συνάρτησης **addedge**.

##### Ερώτημα 4ο

Στην συνέχεια, για κάθε ένα από τα υποδίκτυα για υπολογίζουμε και να παρουσιάσετε γραφικά την κατανομή των τιμών των παρακάτω μέτρων κεντρικότητας:

* 1. Degree Centrality
  2. In-Degree Centrality
  3. Out-Degree Centrality
  4. Closeness Centrality
  5. Betweenness Centrality
  6. Eigenvector Centrality
  7. Katz Centrality

Για τα πρώτα 4 χρησιμοποιήθηκαν οι αντίστοιχες έτοιμες συναρτήσεις του **MATLAB** ενώ για τα τελευταία 2 υλοποιήθηκαν χρησιμοποιώντας τους αντίστοιχους μαθηματικούς ορισμούς τους , καθώς για όλα τα παραπάνω δημιουργήθηκαν τα αντίστοιχα ιστογράμματα.

##### Ερώτημα 5ο

Στο επόμενο ερώτημα , τα σύνολα , αποθηκεύτηκαν στο **cellArray TimeGraphs** ως ένα ενιαίο γράφημα για κάθε χρονικό βήμα, ενώ το σύνολο αποθηκεύτηκε στο **cellArray NewGraphs** ως ξεχωριστό γράφημα σε σχέση με τα προηγούμενα σύνολα για διευκόλυνση του υπολογισμού του πίνακα ομοιότητας **Graph Distance.**

##### Ερώτημα 6ο

Στην συνέχεια για κάθε ζεύγος κόμβων και κάθε σύνολο με υπολογίστηκε με χρήση της ενσωματωμένης συνάρτησης  **distances** ο πίνακας ομοιότητας **Graph Distances** ενώ οι υπόλοιποι ( **common neighbors, jaccard’s coefficient, adamic/adar , preferential attachment** ) με βοήθεια τον γειτόνων ( λόγο κατευθυνόμενου γραφήματος επιλέχτηκαν οι διάδοχοι του κάθε κόμβου) και των μαθηματικών ορισμών τους.

1. [**Common Neighbors**] όπου το σύνολο των γειτόνων του κόμβου .
2. [**Jaccard’s Coefficient**]
3. **[Adamic / Adar]**
4. **[Preferential Attachment]**

##### Ερώτημα 7ο

Τελικά , για κάθε έναν από τους πίνακες ομοιότητας , , και που υπολογίστηκαν στο προηγούμενο ερώτημα και για κάθε ένα από τα σύνολα κόμβων , εξήχθησαν οι κορυφαίες τιμές με βάση τα δεδομένα που έδωσε ο χρήστης στην αρχή του προγράμματος και υπολογίσθηκε πόσες από αυτές υπήρξαν πράγματι στο σύνολο και το ποσοστό επιτυχίας της κάθε μετρικής.

Στον παραδοτέο φάκελο θα βρείτε συνημμένο ένα pdf όπου εμφανίζονται αναλυτικά αποτελέσματα από την εκτέλεση του προγράμματος.

Λόγο σοβαρών τεχνικών προβλημάτων η εκτέλεση του προγράμματος έγινε με μείωση του συνόλου δεδομένων σε 3000 εγγραφές.